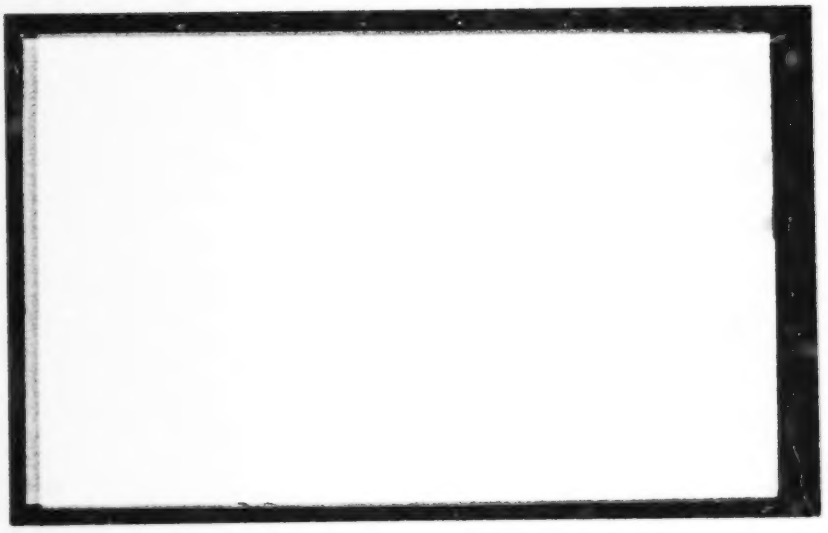
UNIVERSITÉ DE MONTRÉAL



CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

NETWORK DESIGN PROBLEM WITH CONGESTION EFFECTS : A CASE OF BILEVEL PROGRAMMING

by

Patrice Marcotte

Cette recherche a été rendue possible grâce à une subvention du C.R.S.H. (#410-81-0722) et à une subvention du fonds F.C.A.C. (#83-AS-26).

Université de Montréal

Centre de recherche sur les transports - Publication #320
Ecole des Hautes Etudes Commerciales- Cahier du GERAD #G-83-23

Septembre 1983

Dépôt légal - 4e trimestre 1983

Résumé: Récemment plusieurs chercheurs se sont penchés sur des problèmes de planification hiérarchisée qui peuvent être perçus soit comme des jeux de Stackelberg avec variables continues, soit comme une généralisation de problèmes Min-Max. Il est possible de formuler ainsi le problème du design optimal d'un réseau en présence d'effets de congestion. Après avoir décrit formellement le problème, nous proposons divers algorithmes de résolution qui seront étudiés théoriquement et testés numériquement.

Abstract: Recently much attention has been focused on multilevel programming, a branch of mathematical programming that can be viewed either as a generalization of min-max problems or as a particular class of Stackelberg games with continuous variables. The network design problem with continuous decision variables (link capacities) can be cast into such a framework. We first give a formal description of the problem and then develop various heuristic procedures to solve it. Worst-case behaviour results concerning the heuristics, as well as numerical results on a small network, are presented.

1. INTRODUCTION

There are many situations where different levels of decisions are involved. Most of the large-scale mathematical programming literature that deals with decentralized management of resources covers models of such problems where the various decision makers act in a cooperative manner. Models of noncooperative decision making are found in the game theory literature and are usually restricted to a small number of players or variables. Recently some authors have investigated situations where decision makers at various levels act in a hierarchical manner: at a given level, decision makers are bound by the decisions of the upper levels and maximize their own profit accordingly, taking into account the reactions of the lower levels. For a complete and formal description see Bard and Falk [4]. Consider the following bilevel model. If x and y represent the decision vectors associated with the upper and lower level respectively, the problem may be formulated as:

$$\begin{aligned} &\underset{x \in X}{\text{Min}}\ F(x,y) \qquad (1)\\ &\text{subject to: } y = \underset{z \in Y(x)}{\text{argmin}}\ G(x,z) \end{aligned}$$

where X is the feasible set of the x-variables, and Y(x) the feasible set, possible dependent on x, of the y-variables. Algorithms have been developed for linear cost functions (F and G linear, X and Y polyhedral) by Bialas and Karwan [6], Bard and Falk [4], Papavassilopoulos [21], Candler and Townsley [8], and for the general case by Bard [3] and Papavassilopoulos [23].

In this paper we consider a generalization of problem (1) where the y-variables are the solution of an equilibrium problem formulated as a variational inequality. Then the problem becomes:

$$\begin{aligned} &\underset{x \in X}{\text{Min}}\ F(x,y) \qquad (2)\\ &\text{subject to } \quad (y-z)^T H(x,y) \leq 0 \qquad \forall\, z \in Y(x)\\ &\qquad\qquad y \in Y(x) \end{aligned}$$

The network design problem that we consider has already been formulated and studied in a simpler form, by Dafermos [9], Abdulaal and Leblanc [2], Dantzig et al [12] and Gershwin and Tan [26]; it consists in optimally balancing the transportation, investment and maintenance costs of a network subject to congestion, where users behave according to Wardrop's first principle of traffic equilibrium (user-optimum, see Wardrop [27]) which is akin to the Nash-Cournot principle of the theory of games (Friedman [15]). The x-vector represents the improvements made to the network, X the set of feasible improvements, y the multicommodity flow variables and Y(x) the polyhedron of feasible flows, which is independent of x.

The x-variables will reflect the capacities of the network links; for the same link flow, the corresponding link traversal time will be an increasing function of its capacity. In other words: the higher the capacity, the lower the congestion.

The remainder of the paper is structured as follows: first we give a mathematical formulation of the problem then we prove some theoretical results about the problem; finally we propose solution algorithms of a heuristic nature which will be analyzed theoretically and tested numerically.

2. MATHEMATICAL FORMULATION

Throughout this paper, the subscript "a" will refer to a link of the maximal network. Let:

N: set of nodes

A: set of one-way links

r: set of origins

Δ: set of destinations

$v^k = (v_a^k)_{a \in A}$: flow vector of commodity k (trips originating from origin k)

$v_a = \sum_{k \in r} v_a^k$: total flow on link a

$v = (v_a)_{a \in A}$: total flow vector

z_a: capacity of link a

$g_{k\ell}$: trip demand between origin $k \in \Omega$ and destination $\ell \in \Delta$ (element of the O-D matrix).

In the remainder of the paper, we will call a set of flow vectors <u>feasible</u> if it satisfies the following flow conservation and non-negativity constraints:

$$\forall i \in N: \quad \sum_{j \in N} v_{ij}^k - \sum_{j \in N} v_{ji}^k = \begin{cases} \sum_{\ell \in \Delta} g_{k\ell} & \text{if } i = k \\ -g_{ki} & \text{if } i \in \Delta \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

For the sake of brevity we will write $v \in \Phi$ although, formally, we should write $(v_k)_{k \in \Omega} \in \Phi$.

Let $S(v,z)$ be a vector-valued function giving the traversal times of all links as a function of the total flow vector v and the capacity vector $z = (z_a)_{a \in A}$.

It may be shown (Smith [24], Dafermos [10]) that a feasible flow vector v^* constitutes a user equilibrium for a given capacity vector z, if and only if it satisfies the following variational inequality:

$$(v^* - v^{(i)})^T \, S(v^*,z) \leq 0 \qquad i=1,2,\ldots,n \tag{4}$$

for all extremal flows $v^{(i)}$ of Φ .

Whenever the function $S(v^*,z)$ is integrable for fixed z, (4) is equivalent to solving the convex programming problem:

$$\underset{v \in \Phi}{\text{Min}} \int_0^v S(t,z)\, dt \tag{5}$$

The solution of (5) may be efficiently found by using primal methods of feasible directions for minimizing a convex function under linear constraints (see Nguyen [21], Leblanc [17]).

Assumption 1: For z fixed, the function $S(v,z)$ is continuous and strictly monotone increasing for all nonnegative v.

Under assumption 1, solution to (4) exists and is unique (Aashtiani and Magnanti [1]).

Using the previous notations, the bilevel program (2) may be written as:

$$\begin{aligned} \underset{\substack{z\geq 0 \\ v\in\Phi}}{\text{Min}} \quad & v^T S(v,z) + g(z) \\ \text{subject to} \quad & (v - v^{(i)})^T S(v^*,z) \leq 0 \qquad i=1,2,\ldots,N \end{aligned} \tag{6}$$

where $g(z)$ represents the capital investment and operating costs of building and operating a network with capacities $z^{(*)}$.

At the optimum of (6) the Kuhn-Tucker conditions are not necessarily satisfied. Consequently, the algorithm of Bard [3] cannot be applied.

The problem (6) suggests a finitely convergent constraint accumulation algorithm (see Marcotte [20]). This can be contrasted with a somewhat similar result (corollary 2.1) in Blankenship and Falk [7]) where $S(\cdot,z)$ is concave rather than convex.

Assumption 2: The function $g(z)$ is link-separable, i.e.:

$$g(z) = (g_a(z_a))_{a\in A}, \text{ with each function } g_a \text{ increasing.}$$

Also: $g(0) = 0$.

Bilevel problem (2) may be rewritten as:

$$\underset{z\geq 0}{\text{Min}} \quad v(z)^T S(v(z), z) + g(z) \tag{7}$$

where $v(z)$ is the unique solution to variational inequality (4).

Some attempts have been made to solve (6) directly, using descent methods (Leblanc [2]). The difficulty stems from the fact that the objective function, although continuous, is not in general convex or differentiable as a function of the argument z (Dafermos [9], Marcotte [19]). Also, each evaluation of the objective requires the solution of a user equilibrium problem, which is prohibitive.

The next section considers a related problem that is easier to solve and can be used to provide both lower bounds and upper bounds on the true optimum.

(*) Assuming a discount factor $\rho>0$, we have: $g(z)=\frac{\rho}{1+\rho}c(z)+m(z)$ where $c(z)$ is the actual cost of building the z-network and $m(z)$ is the operating cost.

3. SYSTEM-OPTIMUM PROBLEM

If we remove the variational constraint on v, problem (6) becomes:

$$\begin{array}{l} \min_{v\in\Phi} F(v,z) = v^T S(v,z) + g(z) \\ z\geq 0 \end{array} \qquad (8)$$

Let (v^*,z^*) and $(\bar{v},\bar{z})$ be the solutions to (6) and (8) respectively; Since $f(\bar{z})$, the user-optimum flow corresponding to capacity vector $\bar{z}$, is a feasible solution for problem (6), we have the following obvious relations:

$$F(\bar{v},\bar{z}) \leq F(v(z^*),z^*) \leq F(v(\bar{z}),\bar{z}) \ ,$$

Assumption 3: The function $S(v,z)$ is link-separable and has the form:

$S(v,z) = \sum_{a\in A} S_a(\frac{v_a}{z_a})$, where S_a are convex functions.

Proposition 1: The function $v^T S(v,z)$ is convex.

Proof: Let H_a denote the hessian matrix of the function $v_a S_a(\frac{v_a}{z_a})$.

Temporarily dropping subscript a for clarity, we have:

$$H = \begin{bmatrix} \frac{2}{z} S'(\frac{v}{z}) + \frac{v}{z^2} S''(\frac{v}{z}) & -\frac{2v}{z^2} S'(\frac{v}{z}) - \frac{v^2}{z^3} S''(\frac{v}{z}) \\ -\frac{2v}{z^2} S'(\frac{v}{z}) - \frac{v^2}{z^3} S''(\frac{v}{z}) & \frac{2v^2}{z^3} S'(\frac{v}{z}) + \frac{v^3}{z^4} S''(\frac{v}{z}) \end{bmatrix}$$

The determinant of H is:

$$\begin{aligned} \det H &= [\frac{2}{z} S'(\frac{v}{z}) + \frac{v}{z^2} S''(\frac{v}{z})]\cdot[\frac{2v^2}{z^3} S'(\frac{v}{z}) + \frac{v^3}{z^4} S''(\frac{v}{z})] \\ &\quad - [\frac{2v}{z^2} S'(\frac{v}{z}) - \frac{v^2}{z^3} S''(\frac{v}{z})]^2 \\ &= 0 \end{aligned}$$

The trace of H is:

$$\text{tr } H = \frac{2}{z} S'(\frac{v}{z}) + \frac{v}{z^2} S''(\frac{v}{z}) + \frac{2v^2}{z^3} S'(\frac{v}{z}) + \frac{v^3}{z^4} S''(\frac{v}{z})$$

$$\geq 0 \quad \text{since S is increasing and convex.}$$

These results imply that H is positive semidefinite therefore, each $v_a S_a(\frac{v_a}{z_a})$ is convex and $v^T S(v,z) = \sum_{a \in A} v_a S_a(\frac{v_a}{z_a})$ is convex. ∎

Corollary 2. If the function g is convex, then F(v,z) is convex.

Under the hypothesis of corollary 2, problem (8) is thus a convex programming problem in the variables v and z. It can be further simplified (Los[18], Marcotte [20]) by first performing the minimization with respect to the z_a-variables, for each link-flow v_a fixed:

$$\min_{z_a \geq 0} \; v_a S_a(\frac{v_a}{z_a}) + g_a(z_a) \tag{9}$$

whose unique solution will be denoted $z_a(v_a)$. This yields:

$$\min_{v \in \Phi} \; \sum_{a \in A} v_a S_a(\frac{v_a}{z_a(v_a)}) + g_a(z_a(v_a)) \tag{10}$$

This last formulation is equivalent to a standard convex flow assignment problem, for which many efficient algorithms already exist (Nguyen [21], Daganzo [11], Bertsekas and Gafni [5]).

4. LINEAR COST FUNCTION

When the cost function g is linear, both the system-optimum and user-optimum problem possess interesting properties. We first recall a result of Los [18]:

Proposition 3: If the cost function g is linear, i.e.: $g_a(z_a) = \ell_a z_a \ \forall a \in A$ then the objective function in problem (10) is linear.

Proof: By replacing $g_a(z_a)$ by $\ell_a z_a$ in (9) we get the one dimensional minimization problems:

$$\text{Min} \quad v_a S_a\left(\frac{v_a}{z_a}\right) + \ell_a z_a \tag{11}$$

By differentiating with respect to z_a, we obtain the necessary and sufficient condition for a minimum:

$$\left(\frac{v_a}{z_a}\right)^2 S_a'\left(\frac{v_a}{z_a}\right) = \ell_a \tag{12}$$

Let v_a be the solution of (12). Then we have: $z_a = \frac{v_a}{u_a}$. Replacing for z_a in the objective (11) gives:

$$F(v,z) = \sum_{a \in A} (\ell_a u_a + S_a(u_a))v_a$$

which is linear in v. ∎

Proposition 3 shows that the system-optimum problem (11) can be solved by performing shortest path calculations.

The next proposition states that, under slightly more restrictive assumptions, the system-optimum and user-optimum problems are equivalent, notwithstanding the topology of the maximal network nor the values of the entries of the O-D matrix.

Assumption 3: The cost function is linear and each of its components is proportional to the corresponding arc length and capacity, i.e.:

$$g_a(z_a) = \ell d_a z_a$$

where d_a represents the physical length of arc a, and ℓ is a constant.

Assumption 3 implies that the investment and maintenance costs are homogeneous throughout the networks.

Assumption 4: The link traversal times are homogeneous throughout the network and proportional to the arc lengths, i.e.:

$$S_a(\frac{v_a}{z_a}) = d_a\, \sigma(\frac{v_a}{z_a}) \ .$$

Proposition 4: Under assumptions 3 and 4, the solution(s) of the system-optimum problem (8) is feasible and thus optimal for the user-optimum problem (6).

Proof: Under the above assumptions, the system-optimum problem is:

$$\begin{array}{l}\text{Min} \\ v\in\Phi \\ z\geq 0\end{array} \quad v_a d_a\, \sigma(\frac{v_a}{z_a}) + \ell d_a z_a \qquad (13)$$

At the optimum one must have:

$$(\frac{v_a}{z_a})^2\, \sigma'(\frac{v_a}{z_a}) = \ell$$

Let u be the unique solution to the equation:

$$x^2 + \sigma'(x) = \ell \ .$$

We have: $z_a = \frac{v_a}{u}$.

After substitution in (13) we get:

$$\underset{v\in\Phi}{\text{Min}} \sum_{a\in A} d_a\left[\sigma(u) + \frac{\ell}{\psi}\right] v_a$$

which is equivalent to:

$$\underset{v\in\Phi}{\text{Min}} \sum_{a\in A} d_a v_a \,. \tag{14}$$

Let v^* be a (not necessarily unique) optimal solution to (14), and: $z^* = \frac{v^*}{u}$. We must prove that v^* is the user equilibrium corresponding to the capacity vector z^*, i.e. that v^* satisfies the variational inequality:

$$\sum_{a\in A} (v_a^* - v_a^{(i)})\, d_a\, \sigma\left(\frac{v_a^*}{z_a^*}\right) \le 0 \qquad i=1,2,\ldots n. \tag{15}$$

v^* being solution of (14), we have:

$$\sum_{a\in A} (v_a^* - v_a^{(i)})\, d_a \le 0 \qquad i=1,2,\ldots n.$$

Multiplying by the constant $\sigma\left(\frac{v_a^*}{z_a^*}\right) = \sigma(u)$, we get the desired result. ∎

In general however, when assumptions 3 and 4 are not satisfied, characterization of the optimal solution for the user-optimum problem (6) is difficult. The next proposition gives a property possessed by at least one solution flow vector v^*.

Proposition 5: If g is linear, there exists a solution vector (v^*, z^*) to (6) for which v^* is an extremal flow, i.e.: $v^* = v^{(i)}$ for a certain index i.

Proof: We make the change of variable:

$$k_a = \frac{v_a}{\ell_a} \quad .$$

This is equivalent to using the congestion levels on the links of the maximal network as the primary decision variables. The user-optimum problem (6) then becomes:

$$\underset{k \geq 0}{\text{Min}}\ F(v,k) = \sum_{a \in A} v_a S_a(k_a) + \frac{\ell_a v_a}{k_a} v_a \left(S_a(k_a) + \frac{\ell_a}{k_a}\right) \qquad (16)$$

subject to:

$$v \in \{\arg \min_{w \in \Phi} \sum_{a \in A} w_a S_a(k_a)\} \ .$$

Let (v^*, k^*) be a solution to the above problem (16). Since Φ is convex and may be assumed bounded without loss of generality, we can express v^* as a convex combination of extremal points of Φ:

$$v^* = \sum_{i \in I} \lambda_i v^{(i)} \quad \lambda_i > 0 \ , \quad \sum_{i \in I} \lambda_i = 1, \text{ where I is a subset}$$

of $(1,2,\ldots,n)$. Now, v^* is feasible for (6) and must therefore be a minimum for the expression:

$$\sum_{a \in A} w_a S_a(k_a^*)$$

$$\text{i.e.:} \quad \sum_{a \in A} \sum_{i \in I} \lambda_i v_a^{(i)} S_a(k_a^*) \leq \sum_{a \in A} w_a S_a(k_a^*) \quad \forall w \in \Phi$$

$$\Rightarrow \quad \sum_{i \in I} \lambda_i \sum_{a \in A} v_a^{(i)} S_a(k_a^*) \leq \sum_{a \in A} w_a S_a(k_A^*) \quad \forall w \in \Phi \ .$$

The λ_i being positive, we conclude that:

$$\sum_{a \in A} v_a^{(i)} S_a(k_a^*) \leq \sum_{a \in A} w_a S_a(k_a^*) \qquad \begin{matrix} \forall w \in \Phi \\ i \in I \end{matrix}$$

$\Rightarrow$ the $v^{(i)}$'s are feasible for (6), with $k = k^*$.

Now: $$F(v^*,k^*) = \sum_{a\in A}\Big(\sum_{i\in I}\lambda_i v_a^{(i)}\Big)\Big(S_a(k_a^*) + \frac{\ell_a}{k_a^*}\Big)$$

$$= \sum_{i\in I}\lambda_i\left[\sum_{a\in A} v_a^{(i)}\Big(S_a(k_a^*) + \frac{\ell_a}{k_a^*}\Big)\right]$$

$$= F(v^{(i)},k^*) = F(v^*,k^*) \qquad \text{for all } i \text{ in } I$$

$\Rightarrow$ All the $v^{(i)}$'s are optimal. This completes the proof. ∎

Corollary 6. If g is concave, there exists an optimal solution (v^*,z^*) to (6) where v^* is an extremal flow of Φ.

Proof: Using the notations in the proof of proposition 5 we prove, following the same line of reasoning, the feasibility of the extremal flows $v^{(i)}$ for problem (6). Also:

$$F(v^*,k^*) = \sum_{a\in A}\left[\sum_{i\in I}(\lambda_i v_a^{(i)})\, S_a(k_a^*)\right] + \left[g_a\Big(\frac{\sum_{i\in I}\lambda_i v_a^{(i)}}{k_a^*}\Big)\right]$$

$$= F(v^*,k^*) \geq \left[\sum_{i\in I}\lambda_i \sum_{a\in A} v_a^{(i)}\, S_a(k_a^*)\right] + \left[\sum_{i\in I}\lambda_i \sum_{a\in A} g_a\Big(\frac{v_a^{(i)}}{k_a^*}\Big)\right] \qquad (17)$$

by concavity of g_a.

Therefore: $$F(v^*,k^*) \geq \sum_{i\in I}\lambda_i\left[\sum_{a\in A} v_a^{(i)}\, S_a(k_a^*) + g_a\Big(\frac{v_a^{(i)}}{k_a^*}\Big)\right]$$

$$= \sum_{i\in I}\lambda_i F(v^{(i)},k^*)$$

Now: $F(v^*,k^*)$ optimal implies that:

$$F(v^{(i)},k^*) = F(v^*,k^*) \qquad \text{for all} \quad i \text{ in } I$$

$\Rightarrow$ the $v^{(i)}$'s are optimal for (6). ∎

Corollary 7: If g is strictly concave, any solution (v^*,z^*) to (6) must include an extremal flow of v^*.

Proof: In the preceding proof, the inequality (17) is strict(*) unless I is composed of a single element i^*, and:

$$\lambda_{i^*}=1, \qquad v^*=v(i^*) \ . \qquad \blacksquare$$

The last two corollaries are interesting because they do not rest in any way on the concavity (or strict concavity) of the objective function F(v,z). They can be contrasted with the results of Bialas and Karwan [6] pertaining to the linear bilevel programming problem. The following example shows that F may be nonconcave even with the S_a's convex and the g_a's concave.

Let $\quad S_a(\frac{v_a}{z_a}) = \frac{v_a}{z_a}$ and $g_a(z_a) = \text{Log}\ (1 + z_a)$

The hessian matrix H_a of $v_a S_a(\frac{v_a}{z_a}) + g_a(z_a) = \frac{v_a^2}{z_a} + \text{Log}\ (1+z_a)$ is:

$$H_a = \begin{pmatrix} \frac{2}{z_a} & \frac{-2v_a}{z_a^2} \\ \frac{-2v_a}{z_a^2} & \frac{2v_a^2}{z_a^3} - \frac{1}{(1+z_a)^2} \end{pmatrix}$$ whose determinant is:

$\det H_a = \frac{-2}{z_a(1+z_a)^2} < 0$, implying the nonconcavity of the objective f(v,z).

Proposition 5 and its two corollaries, although providing a partial characterization of the optimal solution, cannot be used directly to design a search algorithm. There still remains the problem of optimally associating to an extremal flow $v^{(i)}$ its z-counterpart.

(*) We rule out the rather uninteresting case where all v_a's are zero, i.e. the O-D matrix is empty!

5. HEURISTIC PROCEDURES

The formulation (6) suggests (see Marcotte [20]) the use of a restriction procedure for solving problem (6). Unfortunately the resulting restricted subproblems are nonconvex.

In this section we propose six heuristic procedures whose theoretical and numerical behavior will be analyzed in the next two sections. In the remainder of the paper the congestion functions will be of the BPR-type, i.e.:

$$s_a\left(\frac{v_a}{z_a}\right) = \alpha_a + \beta_a\left(\frac{v_a}{z_a}\right)^p \tag{18}$$

and the cost functions will be power functions: $g_a(z_a)=\delta_a z_a^m$ with $m > 0$.

One heuristic (namely H3) relies on these particular functional forms. The others can be readily applied to the general case.

5.1 Heuristic H0

Let $(\bar{v},\bar{z})$ be solution of the system-optimum problem (8).

Take $(v(\bar{z}), \bar{z})$ as a feasible suboptimal solution.

In view of the difficulty of solving the user-optimum problem, this procedure has been proposed, among others, by Dantzig et al [12].

5.2 Heuristic H1

The motivation behind heuristic H0 was to solve a simpler problem obtained by dropping all variational constraints in (6). In heuristic H1, only the most violated constraint is kept at each iteration, resulting in a mathematical programming problem that is convex except for the additionnal constraint. This leads to the following algorithm:

Algorithm H1

Step 1 $k = -1$ $\quad LB = 0$ $\quad UB = \infty$

Solve the system-optimum problem (8) whose solution is $(\bar{v}, \bar{z})$.

Step 2 $k = k + 1$

Solve the all-or-nothing problem:

$$\min_{v \in \Phi} \; v_a(\alpha_a + \beta_a (\frac{\bar{v}_a}{\bar{z}_a})^p)$$

whose solution is $v^{(1)}$.

Obtain $v(\bar{z}) = \arg\min_{v \in \Phi} \sum_{a \in A} \int_0^{v_a} S_a(\frac{t}{\bar{z}_a})\, dt.$

Set: $UB^{(k)} = F(v(\bar{z}), \bar{z})$

$LB^{(k)} = F(\bar{v}, \bar{z})$

$UB = \text{Min } \{UB, UB^{(k)}\}$

$LB = \text{Max } \{LB, LB^{(k)}\}$

If UB-LB is small, STOP; the best solution corresponds to vectors $v(\bar{z})$ and $\bar{z}$ obtained at the iteration k for which $UB = UB^{(k)}$.

Step 3 Solve the augmented problem:

$$\min_{\substack{v \in \Phi \\ z \geq 0}} \; \sum_{a \in A} v_a(\alpha_a + \beta_a (\frac{v_a}{z_a})^p + \ell_a z_a^m \qquad (19)$$

$$\text{subject to:} \quad \sum_{a \in A} (v_a - v_a^{(1)})(\alpha_a + \beta_a (\frac{v_a}{z_a})^p) \leq 0$$

whose solution is $(\bar{v}, \bar{z})$; return to step 2. Δ

To solve subproblem (19) we will transform it into a nonconvex flow problem in the v-variables only. Consider the function:

$$G(v) = \min_{z \geq 0} \sum_{a \in A} v_a \left(\alpha_a + \beta_a \left(\frac{v_a}{z_a} \right)^p \right) + \ell_a z_a^m \qquad (20)$$

subject to:

$$\sum_{a \in A} (v_a - v_a^{(1)}) \left(\alpha_a + \beta_a \left(\frac{v_a}{z_a} \right)^p \right) \leq 0$$

Proposition 7: The function G(v) is well defined and differentiable for all vectors v different of $v^{(1)}$.

Proof: First we note that the additionnal constraint of problem (19) must be active, being violated at the current iteration.

Rather than working directly on (19), we associate a real Lagrange multiplier with the additionnal constraint to write (19) in the form:

$$G(v) = \max_{\lambda \in R} \min_{z \geq 0} \sum_{a \in A} [v_a + \lambda(v_a^{(1)} - v_a)] [\alpha_a + \beta_a \left(\frac{v_a}{z_a} \right)^p] + \ell_a z_a^m \qquad (21)$$

We perform the minimization with respect to z_a by equating the z_a-partial derivative of G to zero; after some simple algebra we get:

$$z_a = r_a \{ v_a^p [v_a + \lambda(v_a^{(1)} - v_a)] \}^{\frac{1}{m+p}}$$

where $$r_a = \left(\frac{p\beta_a}{m\ell_a} \right)^{\frac{1}{m+p}}$$

Now $\lambda = \lambda(v)$ must be solution of the nonlinear equation:

$$\Gamma(\lambda) = \sum_{a \in A} (v_a - v_a^{(1)}) \left(\alpha_a + \beta_a \left(\frac{v_a}{z_a(v_a, \lambda)} \right)^p \right) = 0 \ . \qquad (22)$$

Replacing z_a by its value we have:

$$\sum_{a \in A} [v_a - v_a^{(1)}] \left[\alpha_a + \frac{\beta_a}{r_a^p} \cdot \left(\frac{v_a^{\frac{1-m}{1+p}}}{1+\lambda\left(\frac{v_a}{v_a^{(1)}} - 1\right)} \right)^p \right] = 0$$

It is easy to see that Γ is a strictly decreasing function of λ with

$$\Gamma\left(\frac{1}{1 - \underset{a \in A}{\text{Max}} \left\{ \frac{v_a^{(1)}}{v_a} \right\}} \right) = +\infty \text{ and } \Gamma\left(\frac{1}{1 - \underset{a \in A}{\text{Min}} \left\{ \frac{v_a^{(1)}}{v_a} \right\}} \right) = -\infty .$$

Thus, for v fixed and distinct of $v^{(1)}$, Γ possesses a unique zero $\lambda(v)$. Differentiability of all functions involved implies the differentiability of $\lambda(v)$ and therefore the differentiability of G(v). ∎

Feasible descent methods may be used to determine local minima of G(v). The partial derivatives $\frac{\partial G(v)}{\partial v_a}$ are obtained by looking for the zero of the derivative of the right hand side of expression (22). Details of the computations will be found in Marcotte [20].

Algorithm cannot be expected to converge at all. It will provide however a sequence of solutions, obtained through a systematic approach.

5.3 Heuristic H2

This procedure has been applied by Gershwin and Tan [26] to the Traffic Signals Setting Problem. One iteration of the algorithm consists in solving sequentially an optimization problem involving the capacity variables, with the flow variables fixed, and then, finding the user-optimized equilibrium corresponding to this new capacity vector. After several iterations of this process, one hopes that convergence toward a "good" solution will be obtained.

Algorithm H2

Step 1 (Initialization). Choose any feasible capacity vector $\bar{z}$.

Step 2 Solve the equilibrium problem:
(equilibrium phase)

$$\underset{v\in\Phi}{\text{Min}} \sum_{a\in A} \int_0^{v_a} S_a\left(\frac{t}{\bar{z}_a}\right)dt \tag{23}$$

whose solution is $\bar{v}$.

Step 3 Solve:
(optimization phase)

$$\underset{z\geq 0}{\text{Min}} \sum_{a\in A} \bar{v}_a S_a\left(\frac{\bar{v}_a}{z_a}\right) + g_a(z_a) \tag{24}$$

whose solution is z^*; set $\bar{z} \leftarrow z^*$.

If $\|\bar{z}-z^*\|<\epsilon$, STOP; otherwise return to step 2. Δ

The optimization phase at step 2 simply consists in finding the roots of the equations:

$$\left(\frac{v_a}{z_a}\right)^2 S'_a\left(\frac{v_a}{z_a}\right) = g'_a(z_a),\ a\in A.$$

It is shown in Marcotte ([19] and [20]) that heuristic H2 is equivalent to solving the following optimization problem:

$$\underset{v\in\Phi}{\text{Min}} \sum_{a\in A} \int_0^{v_a} S_a\left(\frac{t}{z_a(t)}\right)dt \tag{25}$$

where $z_a(v_a)$ is solution of (24).

Under our assumptions about the functional form of S_a and g_a we have:

$$z_a(v_a) = \left(\frac{p\beta_a}{m\ell_a}\right)^{\frac{1}{m+p}} \cdot v_a^{\frac{p+1}{p+m}} . \tag{26}$$

5.3 Heuristic H3

The idea consists in finding a capacity vector for which the system-optimum flow vector $\bar{v}$ is feasible.

Putting the value of $z_a(v_a)$ given by expression into the objective $F(v,z)$ we find:

$$\bar{v} = \arg\min_{v\in\Phi} v_a\left(\alpha_a + d_a v_a^{\frac{p(m-1)}{m+p}}\right) \qquad (27)$$

where $$d_a = (p+m)\left(\frac{\beta_a}{m}\right)^{\frac{m}{m+p}}\left(\frac{\ell_a}{p}\right)^{\frac{p}{m+p}}.$$

In order that the solution be a user-optimized equilibrium, it must satisfy the variational inequality.

$$\sum_{a\in A}(v_a - v_a^{(i)})\left(\alpha_a + \beta_a\left(\frac{v_a}{z_a}\right)^p\right) \le 0 \qquad i=1,\ldots,n \qquad (28)$$

Now, (28) corresponds to the mathematical program:

$$\min_{v\in\Phi}\sum_{a\in A}\int[\alpha_a + \beta_a\left(\frac{t}{\zeta_a(t)}\right)^p]dt \qquad (29)$$

where $\zeta_a(v_a)$ is a function giving the desired value of z_a, and depends on the corresponding link-flow v_a.

(27) and (29) will possess identical solutions if:

$$\int_0^{v_a}[\alpha_a + \beta_a\left(\frac{t}{\zeta_a(t)}\right)^p]\,dt = v_a\left(\alpha_a + d_a v_a^{\frac{p(m-1)}{m+p}}\right)$$

for all $v_a \ge 0$.

This identity is realized if we take:

$$\zeta_a(v_a) = b_a v_a^n \quad \text{with} \quad n = \frac{p+1}{p+m} \quad \text{and} \quad b_a = \frac{1}{(p+1)^{\frac{1}{p}}} \cdot \frac{p\beta_a}{m\ell_a} \; . \tag{30}$$

Thus the solution vector $(\bar{v}, b_a \bar{v}_a^n)$ is feasible for (6).

5.5 Heuristic H4

At each iteration of the Frank-Wolfe algorithm applied to the user-optimized equilibrium problem with link cost functions $S_a(v_a)$, a point is found which is the convex combination of the previous point v^k and of an extremal point y:

$$v^{k+1} = y + \lambda(v^k - y)$$

where $$y \in \{\arg\min_{v \in \Phi} \; v^T S(v^k)\}$$

and λ is chosen in such a way that y cannot be used to generate a descent direction at the next iteration, i.e.:

$$(v^{k+1} - y)^T \, S(v^{k+1}) \leq 0 \tag{31}$$

In heuristic H4, the points generated at each iteration satisfy relation (31), the optimal step λ being a function of the capacity vector z.

Algorithm H4

Step 1 Solve the system-optimum problem (8) with solution $(\bar{v}, \bar{z})$.

Step 2 Find a most violated constraint and the corresponding extremal flow:

$$y \in \{\arg\min_{w \in \Phi} \sum_{a \in A} w_a S_a\left(\frac{\bar{v}_a}{\bar{z}_a}\right)\}$$

Step 3 Define implicitly the function $\lambda(z)$ through the relation:

$$\lambda(z) = \underset{0 \le \mu \le 1}{\arg\min} \sum_{a \in A} \int_{0}^{y_a + \mu(\bar{v}_a - y_a)} S_a(t)\,dt$$

Step 4 Let:
$$\tilde{z} = \underset{z \ge 0}{\arg\min}\ F[y + \lambda(z)(\bar{v}-y),\ z] \qquad (32)$$

if $\|\bar{z}-\tilde{z}\| < \epsilon$ STOP; otherwise return to step 2. Δ

The strict monotonicity of the functions S_a implies that $\lambda(z)$ is well defined and continuous, even at the point y. This was not the case with heuristic H1. Also problem (32) is simpler to solve than problem (19) of algorithm H1, involving only nonnegativity constraints on the capacity vector. When the parameter p of the congestion functions S_a is an integer smaller than 5, $\lambda(z)$ may be found by solving the polynomial equation:

$$\sum_{a \in A} [\bar{v}_a - y_a] \left[\alpha_a + \beta_a \left(\frac{y_a + \lambda(z)(\bar{v}_a - y_a)}{z_a}\right)^p\right] = 0$$

for which exact formulae exist.

For instance, if p=1 and $\lambda^* = -[\sum_{a \in A} (\bar{v}_a - y_a)(\alpha_a + \frac{\beta_a y_a}{z_a})\}/[\sum_{a \in A} \frac{(\bar{v}_a - y_a)^2}{z_a}]$

we have:
$$\lambda(z) = \begin{cases} 0 & \text{if } \lambda^* \le 0 \\ \lambda^* & \text{if } 0 \le \lambda^* \le 1 \\ 1 & \text{if } \lambda^* \ge 1 \end{cases}$$

For general congestion functions (not necessarily of the BPR-type) the functions $z_a(v_a)$ in heuristic H1 cannot be obtained in closed form; this increases considerably the computational burden, making the objective in (21) a "two-level" implicit function $F(\lambda(z(v)))$. This does not occur in heuristic H4 where, at worst, only $\lambda(z)$ need to be computed numerically.

The drawbacks of H4 with respect to H1 are twofold:

1. Even if it converges, H4 does not necessarily yield the optimum.

2. At step 4 of the algorithm, $F[y + \lambda(\tilde{z})(\bar{v}-y), \tilde{z}]$ is not necessarily a lower bound for the optimum $F(v^*, z^*)$.

5.6 Heuristic H5

Heuristic H2 and H3 may be considered as particular cases of a general procedure where we find feasible solutions to (6) by solving a convex problem from which the variational constraints are absent. Let $J(v,z)$ be the objective of this modified problem, where:

$$J(v,z) = \sum_{a\in A} \int_0^{v_a} [\alpha_a + \beta_a (\frac{t}{z_a})^p] dt + \xi_a \ell_a z_a^m \qquad (33)$$

Heuristic H2 corresponds to the choice: $\xi_a = (p+1)^m$ for all links.

The problem of selecting optimal ξ_a's is of course equivalent to the original problem. We can however make educated guesses, the simplest being to set all ξ_a's to a given value ξ. This gives the objective function for heuristic H5:

$$\underset{\substack{v\in\Phi \\ z\geq 0}}{\text{Min}} \sum_{a\in A} v_a [\alpha_a + \frac{\beta_a}{p+1} (\frac{v_a}{z_a})^p] + \xi \ell_a z_a^m \qquad (34)$$

This corresponds to minimizing the weighted sum of user-perceived travelling times and of the investment and maintenance costs of the network.

6. WORST-CASE ANALYSIS OF THE HEURISTICS

In this section we develop upper bounds on the ratio of the values of heuristic and exact solutions. The analysis does not depend on the network structure. Heuristics H1 and H4 are not part of this study, and we can only state that their behaviour will be at least as good as that of heuristic H0.

Definition:
$$R_m^p(H) = \sup_{\substack{\alpha,\beta,\ell \\ N,A}} \frac{\text{Cost of heuristic solution}}{\text{Cost of optimal solution}} .$$

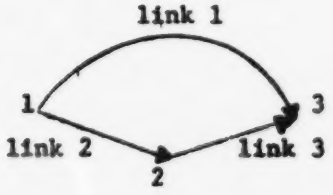
$$= \sup_{\substack{\alpha,\beta,\ell \\ N,A}} \frac{F(v(z^H), z^H)}{F(v^*, z^*)}$$

where z^H is the capacity vector obtained from heuristic H (H = H0, H2, H3 or H5).

6.1 Heuristic H0

Proposition 8: $\lim_{p\to\infty} R_1^p(H0) \geq 2.$

Proof: Consider the network of figure 1.



$$s_{12} = s_{23} = \Pi = p$$

$$s_{13} = 1$$

Figure 1.

We set: $S_1(x) = p$ $\quad S_2(x) = S_3(x) = \frac{1}{2}(1+x^{p^2})$

$g_1(x) = 0$ $\quad g_2(x) = g_3(x) = \frac{1}{2} p\, x$

Since links 2 and 3 play similar roles, we can replace them by a fictitious link 4 that must carry a flow of at least Π units, and with congestion and cost functions given by:

$$S_4(x) = 1 + x^{p^2} \qquad g_4(x) = px\ .$$

The system-optimum problem is:

$$\begin{array}{l}\text{Min} \\ v\in\Phi \\ z\geq 0\end{array} \quad p\, v_1 + v_4\left[1 + \left(\frac{v_4}{z_4}\right)^{p^2}\right] + p\, z_4$$

The optimum z_4 must satisfy:

$$p^2 \left(\frac{v_4}{z_4}\right)^{p^2+1} = p$$

and the problem becomes:

$$\begin{array}{l}\text{Min} \\ v\in\Phi\end{array} \quad pv_1 + v_4\left[1 + \left(\frac{1}{p}\right)^{\frac{p^2}{p^2+1}} + p\ .\ \frac{p^2+2}{p^2+1}\right]$$

The optimal solution is: $\bar{v}_1 = 1$

$$\bar{v}_4 = p \qquad \bar{z}_4 = p^{\frac{p^2+2}{p^2+1}}$$

The cost of this solution is: $F(\bar{v},\bar{z}) = p^2 + 0(p)$.

Now the equilibrium is not reached since travel time on link 4 is less than the corresponding travel time on link 1. A fraction of travellers will therefore switch from link 1 to link 4, until travel times on both links are equal to p. This equilibrium occurs before v_1 gets to zero, since we have:

$$1 + \left(\frac{\bar{v}_4+1}{\bar{z}_4}\right)^{p^2} = 1 + \frac{p+1}{p^{\frac{p^2+2}{p^2+1}}} > p \quad \text{for} \quad p \text{ sufficiently large.}$$

The cost of this solution is:

$$F(v(\bar{z}),\bar{z}) = 2p^2 + 0(p) .$$

The optimal solution consists in slightly increasing the capacities on arc 4 to cater for all (p+1) travellers; the corresponding cost is:

$$F(v^*,z^*) = p^2 + 0(p)$$

$$\Rightarrow \lim_{p\to\infty} \frac{F(v(\bar{z}),\bar{z})}{F(v^*,z^*)} \geq 2 . \qquad \blacksquare$$

6.2 Heuristic H2

Proposition 9: R_1^p (H2) = p+1.

Proof: For m=1 and $S_a(x) = \alpha_a + \beta_a x^p$, problem (25) takes the form:

$$\underset{v\in\Phi}{\text{Min}} \sum_{a\in A} v_a\left[\alpha_a + \beta_a \left(\frac{\ell_a}{p\beta_a}\right)^{\frac{p}{p+1}}\right] \tag{35}$$

Let $\tilde{v}$ be solution of (35) and $\tilde{z}_a = z_a(\tilde{v}) = \left(\frac{p\beta_a}{\ell_a}\right)^{\frac{1}{p+1}} \tilde{v}_a$.
We have:

$$F(\tilde{v},\tilde{z}) = \sum_{a\in A} \tilde{v}_a\left[\alpha_a + \beta_a \left(\frac{\tilde{v}_a}{\tilde{z}_a}\right)^p\right] + \ell_a \tilde{z}_a$$

$$= \sum_{a\in A} \tilde{v}_a\left[\alpha_a + (p+1)\,\beta_a \left(\frac{\ell_a}{p\beta_a}\right)^{\frac{1}{p+1}}\right]$$

$$\leq (p+1) \sum_{a\in A} \tilde{v}_a[\alpha_a + \beta_a(\frac{\ell_a}{p\beta_a})^{\frac{p}{p+1}}]$$

$$\leq (p+1) \sum_{a\in A} \bar{v}_a[\alpha_a + \beta_a(\frac{\ell_a}{p\beta_a})^{\frac{p}{p+1}}] = (p+1)F(\bar{v},\bar{z})$$

where $\bar{v}$ is solution of the system-optimum problem.

Finally: $F(\tilde{v},\tilde{z}) \leq (p+1)F(v,z) \leq (p+1)\ F(v^*,z^*)$

This proves that $R_1^p(H2) \leq p+1$. To prove the reverse inequality, consider the network of figure 2:

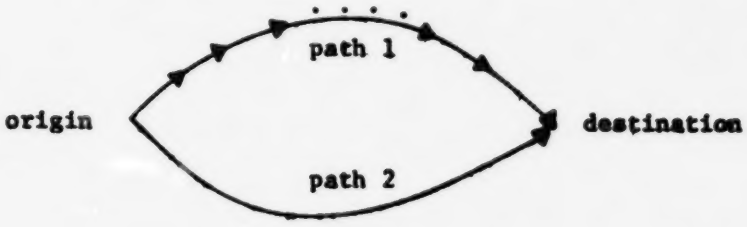


Figure 2.

Path 1 is composed of n identical arcs for which $\alpha_a=\beta_a=1$ and $\ell_a=p(\frac{1}{n})^{\frac{p+1}{p}}$.

Path 2 comprises a unique arc with $\alpha_a=\beta_a=1$ and $\ell_a=p(n-1)^{\frac{p+1}{p}}$.

One unit of flow must travel from the orignl to the destination.

Heuristic H2 yields the path flows;

$\tilde{v}_1 = 0$ $\tilde{v}_2=1$. with total cost

$F(\tilde{v},\tilde{z}) = 1 + (p+1)n.$

The optimal solution consists in routing all the flow on path 1, with costs: $F(v^*,z^*) = n+p+1$.

Therefore: $\lim_{n\to\infty} \frac{F(\tilde{v},\tilde{z})}{F(v^*,z^*)} = \lim_{n\to\infty} \frac{n(p+1)+1}{n+p+1} = p+1$. ∎

6.3 Heuristic H3

Proposition 10: $R_m^p(H3) = \frac{m(p+1)}{m+p} + \frac{p}{(m+p)(p+1)^{m/p}}$.

Proof: Let $(\bar{v},\bar{z})$ be the solution given by heuristic H3. From (30) we find its cost:

$$F(\bar{v},\bar{z}) = \sum_{a\in A} v_a\{\alpha_a + [\frac{m(p+1)}{p+m} d_a + \ell_a(\frac{\beta_a}{d_a}\cdot\frac{m+p}{m(p+1)})^{\frac{m}{p}}] v_a^{\frac{p(m-1)}{p+m}}\}$$

The error is maximum when all α_a are zero; this yields:

$$R_m^p(H3) = \sup_{\beta_a,\ell_a,d_a} \frac{m(p+1)}{m+p} + [\frac{m+p}{m(p+1)}]^{\frac{m}{p}} \frac{\beta_a \ell_a}{d_a^{\frac{m+p}{p}}}$$

$$= \frac{m(p+1)}{m+p} + \frac{p}{(m+p)(p+1)^{\frac{m}{p}}}$$ ∎

For m=1 we get:

$$R_1^p(H3) = 1 + \frac{p}{(p+1)^{1+\frac{1}{p}}} \le p+1 = R_1^p(H2)$$

Corollary 11: $\lim_{\substack{m\to 0\\ m>0}} R_m^p(H3) = 1$ and $\lim_{p\to\infty} R_m^p(H3) = 2$.

Proof: The proof is straightforward. ∎

The first statement of Corollary 11 indicates that, in the important particular case where the investment costs are strongly concave (economies of scale), the solutions to the user-optimised and system-optimised problems will be almost similar. Concavity of the g_a's leads however to a concave optimization problem for which no efficient solution procedures exist yet. The limit case m=0 corresponds to the discrete network design problem (without congestion) and has been treated extensively in the literature: Wong [28], Hoang [16], Dionne [13], etc.

6.4 Heuristic H5

The following lemma will be used to find bounds for $R_1^p(\xi)$.

Lemma 12. Let $\gamma = (\gamma_a)_{a\in A}$ and $\delta = (\delta_a)_{a\in A}$ be two arc length vectors of non-negative numbers and k>0. Let:

$$\bar{v}\in\{\arg\min_{v\in\Phi} \bar{F}(v) = (\alpha+\beta)^T v\} \tag{36}$$

$$\bar{\bar{v}}\in\{\arg\min_{v\in\Phi} \bar{\bar{F}}(v) = (\alpha+k\beta)^T v\} \quad . \tag{37}$$

Then we have: $$\sup_{\alpha,\beta,\Phi} \frac{\bar{\bar{F}}(\bar{v})}{\bar{\bar{F}}(\bar{\bar{v}})} = \text{Max}\ \{k,\frac{1}{k}\} \ .$$

Proof: Suppose k is greater than 1 and consider a network with two parallel arcs linking an origin and a destination. Let $\alpha=(n,1)$, $\beta=(0,n-1-\frac{1}{n})$ for n large. We have: $\bar{v} = (0,1)$ since

$$\alpha_1 + \beta_1 = n > 1 + n-1 - \frac{1}{n} = \alpha_2 + \beta_2 \ .$$

Also: $\bar{\bar{v}} = (1,0)$ since

$$\alpha_1 + k\beta_1 = n < 1 + k(n-1-\frac{1}{n}) = \alpha_2 + k\beta_2 \ ,$$

for n sufficiently large. Hence:

$$\frac{\bar{\bar{F}}(\bar{v})}{\bar{F}(\bar{v})} = \frac{1\ k(n-1-\frac{1}{n})}{n} \xrightarrow[n\to\infty]{} k \quad .$$

$$\Rightarrow \sup_{\alpha,\beta,\Phi} \frac{\bar{\bar{F}}(\bar{v})}{\bar{F}(\bar{v})} \geq k \ .$$

But: $\bar{\bar{F}}(\bar{v}) \leq k\ \bar{F}(\bar{v}) \leq k\ \bar{F}(\bar{v}) \leq k\ \bar{\bar{F}}(\bar{v})$

$$\Rightarrow \sup_{\alpha,\beta,\Phi} \frac{\bar{\bar{F}}(\bar{v})}{\bar{F}(\bar{v})} \leq k \Rightarrow \sup_{\alpha,\beta,\Phi} \frac{\bar{\bar{F}}(\bar{v})}{\bar{F}(\bar{v})} = k \ .$$

If k is less than 1 we replace problem (37) by the equivalent problem:

$\min_{v\in\Phi} (\frac{\alpha}{k} + \beta)^T v$. Reasoning exactly as before, exchanging the roles of α and β, we find:

$$\sup_{\alpha,\beta,\Phi} \frac{\bar{\bar{F}}(\bar{v})}{\bar{F}(\bar{v})} = \frac{1}{k} \quad .$$

Removing the condition on k yields the desired result. ∎

Proposition 13:

$$1 + \frac{p}{\xi(p+1)} \leq R_1^p(H5) \leq \frac{\xi^{\frac{p}{p+1}}}{(p+1)^{\frac{1}{p+1}}} \left[1 + \frac{p}{\xi(p+1)}\right]^2 \ .$$

Proof: After substituting the optimal value of z_a into the objective of (33) and setting m to 1, we get the problem:

$$\min_{v\in\Phi} \sum_{a\in A} v_a \left[\alpha_a + \frac{\xi^{\frac{p}{p+1}}}{(p+1)^{\frac{1}{p+1}}} (p+1)\ \beta_a^{\frac{1}{p+1}}\ \ell_a^{\frac{p}{p+1}}\right] \quad . \tag{38}$$

The real cost of the solution v^{ξ} to (33) is:

$$F(v^{(\xi)},z^{(\xi)}) = \sum_{a\in A} v_a^{(\xi)} \left[\alpha_a + \beta_a \left(\frac{v_a^{(\xi)}}{z_a^{(\xi)}}\right)^p\right] + \ell_a z_a^{(\xi)}$$

where $z_a^{(\xi)} = \left[\frac{p\,\beta_a}{(p+1)\xi\ell_a}\right]^{\frac{1}{p+1}} v_a^{(\xi)}$, for all a in A.

After replacing $z_a^{(\xi)}$ by its value we obtain:

$$F(v^{(\xi)}, z^{(\xi)}) = \sum_{a\in A} v_a^{(\xi)}\left[\alpha_a + \frac{1}{p+1}\left(\xi^{\frac{p}{p+1}}(p+1)^{\frac{1}{p+1}} + \frac{p}{\xi^{\frac{1}{p+1}}(p+1)^{\frac{1}{p+1}}}\right) e_a\right] \qquad (39)$$

with $e_a = (p+1)\beta_a^{\frac{1}{p+1}} \ell_a^{\frac{p}{p+1}}$.

Applying lemma 12 to expressions (38) and (39) we get:

$$R_1^p(\xi) \geq \frac{1}{p+1}\ \xi^{\frac{p}{p+1}}(p+1)^{\frac{p}{p+1}} + \frac{p}{\xi^{\frac{1}{p+1}}(p+1)^{\frac{1}{p+1}}} \cdot \frac{(p+1)^{\frac{1}{p+1}}}{\xi^{\frac{p}{p+1}}}$$

$$= \frac{1}{p+1}(p+1+\frac{p}{\xi}) = 1 + \frac{p}{\xi(p+1)} > 1 \ .$$

This proves the first inequality. Now consider the system-optimum cost function, obtained after replacing z by its value:

$$F(v,z(v)) = \sum_{a\in A} v_a(\alpha_a + e_a) \qquad (40)$$

Applying again lemma 12 to (38) and (40) the second inequality will hold if the quantity $\sigma = \frac{\xi^{\frac{p}{p+1}}}{(p+1)^{\frac{1}{p+1}}}\ (1 + \frac{p}{\xi(p+1)})$ is greater or equal to 1.

Letting $\frac{p}{p+1} = q$ we have:

$$\sigma(\xi) = \frac{1}{\left(\frac{1}{1-q}\right)} \; \xi^q \left(1 + \frac{q}{\xi}\right)$$

The minimum value of $\sigma(\xi)$ is obtained when:

$$q\,\xi^{q-1} + q(q-1)\,\xi^{q-2} = 0$$

$$\Leftrightarrow \quad \xi = 1-q$$

and: $$\sigma(\xi) \geq (1-q)^{1-q}\,(1-q)^q\,\left(1 + \frac{q}{1-q}\right)$$

$$= (1-q)^{1-q}\,\frac{(1-q)^q}{1-q} = (1-q)^{1-q}\,(1-q)^{q-1} = 1 \qquad \blacksquare$$

<u>Corollary 14</u>: For $\xi=1$ we have: $2 \leq \lim_{p\to\infty} R_1^p(H5) = 4$.

<u>Proof</u>: Replace ξ by 1 in proposition 13; we have:

$$\lim_{p\to\infty} \; 1 + \frac{p}{p+1} = 2 \quad \text{and:}$$

$$\lim_{p\to\infty} \frac{1}{\frac{1}{(p+1)^{\frac{1}{p+1}}}} \left(1 + \frac{p}{p+1}\right)^2 = 4. \qquad \blacksquare$$

In order to obtain strong bounds for heuristic H5, we can choose $\xi = \xi(p)$ such as to minimize the term $U = \sigma\left(1 + \frac{p}{\xi^{(p+1)}}\right)$. We make the change of variable: $\frac{p}{p+1} = q$, and we rewrite:

$$U(\xi) = (1-q)^{1-q}\,\xi^q\left(1 + \frac{q}{\xi}\right)^2 .$$

We want to minimize the expression $V(\xi) = \xi^q(1 + \frac{q}{\xi})^2$ with respect to the variable ξ.

$$V(\xi) = \xi^q + 2q\xi^{q-1} + q^2\xi^{q-2}$$

The minimum value is obtained at the zero of the derivative:

$$V'(\xi) = \xi^{q-3}(\xi^2 + 2(q-1)\,\xi + q(q-2)) = 0$$

$$\Rightarrow \quad \xi = 2-q = 2 - \frac{p}{p+1} \cdot \frac{p+2}{p+1} \quad .$$

Remarks: For $\xi = \frac{1}{p+1}$ (heuristic H2), we have:

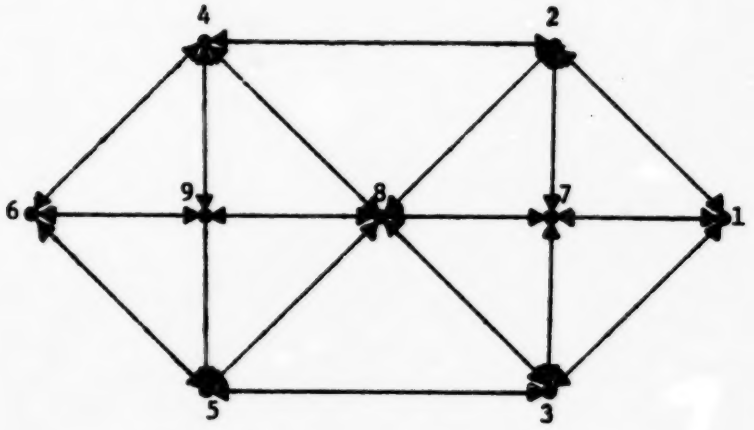
$\sigma = 1$, and: $R_1^p(H2) = p+1$, which corresponds to the statement of proposition 9.

For $\xi = p+1$ (heuristic H3) we have:

$R_1^p(H5) \geq 1 + \frac{p}{(p+1)^2}$ and is exact for $p=1$. For p less than 1, it provides a valid upper bound; for p greater than 1 it is not valid anymore, but the upper bound is much too large, being $O(p)$. In general the upper bound on $R_1^p(H5)$ will be quite pessimistic.

7. COMPUTATIONAL RESULTS

The numerical tests have been performed on a network taken from Steenbrink [25] and reproduced in Figure 3. It is characterized by a large number of paths between pairs of nodes, and these paths share many common arcs. The first data set utilizes Steenbrink's congestion functions. The second data set utilizes a much denser O-D matrix; parameters have been chosen in such a way as to induce high congestion levels, hoping that there would be significant differences between the system-optimum and user-optimum solutions. However the differences were so slight that algorithms H1 and H4 could not improve the solution provided by heuristic H0, and have not been run consequently. In one instance the H2-solution was actually better than the system-optimum solution! This result can be explained by numerical inaccuracies arising while solving convex flow problems. These were solved using a modification by Florian [14] of the Frank-Wolfe requiring more computer memory, but with much better convergence properties. Further details on these tests may be found in Marcotte [20].



9 nodes
36 one-way links

Figure 3.

Tables 1 and 2 give the ratios of the heuristic solutions to the system-optimum solution, for data sets 1 and 2 respectively.

	p	m	H0	H2	H3	H5(ξ=1)
1	1	1	1.003	1.003	1.080	1.023
2	1	2	1.002	1.001	1.202	1.025
3	1	3	1.018	1.034	1.270	1.038
4	2	1	1.000	1.000	1.202	1.082
5	2	2	1.002	1.001	1.583	1.134
6	4	1	1.000	1.000	1.379	1.216
7	4	2	1.000	.999*	1.953	1.358

*due to imperfect convergence

Table 1. (data set 1)

In the first experience with data set 1, the values of the H0 and H2 solutions are identical, although the flows and capacities are quire different. This suggests the existence of distinct optima.

	p	m	H0	H2	H3	H5(ξ=1)
1	1	1	1.000	1.019	1.193	1.068
2	1	2	1.002	1.005	1.367	1.046
3	2	1	1.003	1.006	1.320	1.139
4	2	2	1.001	1.003	1.627	1.146
5	4	1	1.002	1.007	1.449	1.256
6	4	2	1.000	1.001	1.936	1.350

Table 2. (data set 2)

Solutions given by heuristic H0 are consistently the best, and so close to the system-optimum value, that they are, to all extent, optimal.

This good performance of H0 had been previously reported in Los [28]. One should note finally that, contrary to heuristics H3 and H5, the quality of the H0 and H2 solutions does not seem to be much influenced by variations of the parameters p and m.

8. CONCLUSION

In this paper we have conducted a fairly extensive study of a particular kind of continuous, nonlinear network design problem, which can be modelled as a bilevel programming problem. From this study we can draw the following conclusions:

8.1 There is not much hope of developping solution algorithms for finding the global optimum, at least in the general case, and for networks of reasonnable size.

8.2 When the congestion functions are of the BPR-type, and the cost functions are power functions, heuristics H0 and H2 give quasi-optimal solution. This result does not hold however if only some links of the base network are subject to improvement.

8.3 The overall complexity of heuristics H1 and H4 more than offsets the eventual improvements to solutions provided by either H0 or H2.

REFERENCES

[1] H.A. Aashtiani and T.L. Magnanti, "Equilibria on a congested transportation network", *SIAM Journal on Algebraic and Discrete Methods* 2 (1981).

[2] M. Abdulaal and L.J. Leblanc, "Continuous equilibrium network design models", *Transportation Research* B, 13B#1 (1979).

[3] J.F. Bard, "An algorithm for solving the general bilevel programming problem", *Mathematics of Operations Research* 8 #2 (1983).

[4] J.F. Bard and J.E. Falk, "An explicit solution to the multilevel programming problem", *Computers and Operations Research* 9 (1982).

[5] D. Bertsekas and E.M. Gafni, "Projection methods for variational inequalities with application to the traffic assignment problem", *Mathematical Programming Study* 17 (1982).

[6] W.F. Bialas and M.H. Karwan, "On two-level optimization", *IEEE Trans. Automatic Control* AC-27 (1982).

[7] J.W. Blankenship and J.E. Falk, "Infinitely constrained optimization problems", *J. Optimisation Theory and Applications* 19 #2 (1976).

[8] W. Candler and R.J. Townsley, "A linear two-level programming problem", *Computers and Operations Research* 9 (1982).

[9] S.C. Dafermos, "Traffic assignment and resource allocation in transportation networks", Ph.D. Dissertation, John Hopkins University (Baltimore, Maryland 1968).

[10] S.C. Dafermos, "Traffic equilibrium and variational inequalities", *Transportation Science* 14 (1980).

[11] C. Daganzo, "Stochastic network equilibrium with multiple vehicle types and asymmetric, indefinite link cost Jacobians", to appear in *Transportation Science*.

[12] G.B. Dantzig et al, "Formulating and solving the network design problem by decomposition", *Transportation Research* B, 13B#1 (1979).

[13] R. Dionne and M. Florian, "Exact and approximate algorithms for optimal network design", *Networks* 9 (1979).

[14] M. Florian, "An improved linear approximation algorithm for the network equilibrium (Packet Switching) problem", *Proc. of the IEEE Conference on Decision and Control* (1977).

[15] J.W. Friedman, "Oligopoly and the theory of games", North-Holland, Amsterdam (1977).

[16] H.H. Hoang, "A computational approach to the selection of an optimal network, *Management Science* 19 (1973).

[17] L.J. Leblanc, "Mathematical programming algorithms for large scale network equilibrium and network design problems", Ph.D. Dissertation, Northwestern University (1973).

[18] M. Los, "A discrete-convex programming approach to the simultaneous optimization of land-use and transportation", *Transportation Research* 13B (1979).

[19] P. Marcotte, "Network optimization with continuous control parameters", *Transportation Science* 17 #2 (1983).

[20] P. Marcotte, "Design optimal d'un réseau de transport en présence d'effets de congestion", Ph.D. Thesis, University of Montréal (Canada 1981).

[21] S. Nguyen, "An algorithm for the traffic assignement problem", *Transportation Science* 8 (1974).

[22] G. Papavassilopoulos, "Algorithms for leader-follower games", *Proc. of the 18th Annual Allerton Conference on Communication Control and Computing* (1980).

[23] G. Papavassilopoulos, "Algorithms for static Stackelberg games with linear costs and polyhedral constraints", *Proc. of the 21st IEEE Conference on Decision and Control* (1982).

[24] M.J. Smith, "The existence, uniqueness and stability of traffic equilibrium", *Transportation Research* 13B #4 (1979).

[25] P.A. Steenbrink, "Optimization of transportation networks", Wiley, (New-York 1974).

[26] H.N. Tan, S.B. Gershwin and M. Athans, "Hybrid optimization in urban traffic networks", *MIT Report* DOT-TSC-RSPA-79-7 (1979).

[27] J.G. Wardrop, "Some theoretical aspects of road traffic research", *Proc. Inst. Civil Engineers*, Part II, 1 (1952).

[28] R.T. Wong, "A survey of network design problems", *MIT Operations Research Center*, paper #OR80-78 (1978).